

## DTN 中基于泛模板运算的运动模式发现机制

周晓波 张 幸 彭 敏 卢汉成 洪佩琳  
(中国科学技术大学电子工程与信息科学系 合肥 230027)

**摘 要:** DTN(Delay-Tolerant Network) 用于描述 Ad hoc 等无线网络中频繁发生长时间网络分割情形。DTN 不假设端到端路径一定存在, 因此其主要关注点是如何提高成功投递率, 而不是延时等 QoS 参数。运动模式, 例如聚类性等对 DTN 很重要; 本文关注于一种宏观的节点运动模式, 并提出一个发现和使用运动模式的框架——TOM<sup>2</sup>D(Template-Operation based Mobility Model Discovery)。其基本思想是: 节点利用路由信息交互机制维护一个包含所有节点(或相关节点)三维链路容量矩阵, 利用图像处理中的模板运算机制从中提取出可能的运动模式, 最后用一个通用的数据结构存储这些信息并用于路由决策过程中。由于 TOM<sup>2</sup>D 并不依赖于特定的路由协议, 本文最后给出了一个基于 OLSR 和 DSDV 的例子, 仿真结果显示 TOM<sup>2</sup>D 确实为路由决策给出了有价值的信息, 提高了路由成功率。

**关键词:** DTN; TOM<sup>2</sup>D; 运动模式; 宏移动; 模板运算

**中图分类号:** TP393

**文献标识码:** A

**文章编号:** 1009-5896(2009)02-0472-04

## General Template-Operation Based Mobility Model Discovery Mechanism in DTN

Zhou Xiao-bo Zhang Xing Peng Min Lu Han-cheng Hong Pei-lin  
(Dept. EEIS., University of Science and Technology of China, Hefei 230027, China)

**Abstract:** DTN(Delay-Tolerant Network) describes the situation in which longtime partition often happens. DTN doesn't assume the existence of End-to-End path, so it focuses on successful delivery ratio of packets. Nodes' mobility model such as group, is an attractive field in DTN research; this paper concentrates on Macro-Mobility and presents a framework to detect and use mobility models—TOM<sup>2</sup>D(Template-Operation based Mobility Model Discovery). The main idea is: each node maintains a 3D matrix which contains the capacities of links through routing information exchange, then detects the mobility models from the matrix by template-operations, and stores them in a universal data structure at last. This paper gives an example of using TOM<sup>2</sup>D in OLSR and DSDV, and the results of simulation show that TOM<sup>2</sup>D performs well.

**Key words:** DTN; TOM<sup>2</sup>D; Mobility model; Macro-mobility; Template-operation

### 1 引言

DTN(Delay-Tolerant Network)是一种描述无线网络环境, 如 Ad hoc 中频繁发生长时间网络分割情形的网络模型。传统 Ad hoc 网路虽然以考虑网络拓扑变化为重要研究内容, 但是隐含了一个假设: 在路由期间, 从源端到目的端存在一条路径。很多的路由协议如 AODV<sup>[1]</sup>、DSR<sup>[1]</sup>等都依赖于这个假设; DSDV<sup>[1]</sup>和 OLSR<sup>[1]</sup>类的表驱动协议虽然不受这个限制, 但其性能受到很大影响。为克服这一点, 人们开始把节点行为特点作为一个考察对象, 即运动模式, 希望从中提取更多的路由 Metric。典型的方案有: 研究节点的聚合特性, 对节点进行分组, 从而预测发生网络分割的边界<sup>[2]</sup>, 但文献

[2]并没有关注预测结果的使用; 研究节点访问地点信息, 建立 Markov 链模型, 从概率的角度预测节点下一时刻出现在何处<sup>[3,4]</sup>, 并以此作为 Metric; 研究节点间的交互性, 定义“交互矩阵”(IM: Interactive-Matrix), 路由决策转化为对 IM 的处理<sup>[5]</sup>; 研究节点的社会模型<sup>[6]</sup>。节点的运动事实上体现了其持有者的运动, 而人的运动具有很强的规律性, 即文献[6]所引入的“宏移动”和“社会轨迹”的概念——例如对学校里的笔记本而言, 学生会会在周一到周五的早上把它带到图书馆, 下午带到实验室, 晚上带回寝室; 这种周期特性就是社会轨迹。

本文是受文献[6]的启发, 并把社会轨迹的概念一般化, 提出宏观运动模式的概念。就已有的关于运动模式的研究来看, 基本思路都是通过对节点的运动参数, 如速度、方向、关联度和访问点列表的分析, 提炼出一个数值用于路径查找过程。但是并不是所有的运动模式都能简单的用一个数值来

2007-09-13 收到, 2008-04-03 改回

国家自然科学基金青年基金(60602018, 60832005), 微软基金(07122801), 安徽省自然科学基金(070412048)和国家自然科学基金(60772033)资助课题

表达,例如:一个节点周期性的访问一些节点,或者一些节点定期聚合到特定的地方等。因此,本文用一个数据结构 CL(Contact List)作为基本元素,任何运动模式都使用 CL 或 CL 的组合来描述;而且,CL 也将作为路由算法的一个操作对象参与路由决策。另一方面,节点的运动模式所蕴含的时间信息在已有的研究中没有被关注,即便是对社会轨迹的研究,最终也是归结到一个简单的 Markov 概率模型。而时间特性是运动模式的一个维度,例如周期运动模式必须有该节点和每个节点的连接时间信息才能完整表达。而且,时间粒度对运动特征的提取有重要的作用,对于前面提到的校园学生的例子来说,以小时为单位来看是具有一定的周期性,如果以天或分钟来看则其运动可能没有任何特点。TOM<sup>2</sup>D 提供一种可调粒度的运动模式提取方案。

本文以下篇幅将详细介绍 TOM<sup>2</sup>D,首先在第2节描述 TOM<sup>2</sup>D 的流程,然后在第3节把 TOM<sup>2</sup>D 用于 DSDV 和 OLSR,并对其进行仿真,最后在第4节进行总结。

## 2 TOM<sup>2</sup>D 框架

### 2.1 TOM<sup>2</sup>D 的基本步骤

TOM<sup>2</sup>D 中的运动模式是基于对链路容量的分析,而不是对节点运动的物理参数(如位置、速度和方向等)本身。因为一方面,数据的投递过程取决于链路容量,诸如位置、速度之类的物理量可能只是表象,依靠这些参量识别出来的运动模式可能无法使用。例如在物理空间上两个离的很近的节点可能由于干扰问题无法通信,这样它们本质上距离很远。其次,物理参数的获取很困难,可能需要特殊的底层技术甚至外界的辅助(例如 GPS)。最后,因为不同的物理量有不同含义,即有不同的量纲,不能直接运算,因此不能用统一的方式进行处理。而 TOM<sup>2</sup>D 避开了这些问题,关注于链路容量的变化(本文只关注二值化的链路容量,即只有通和断两个状态),提取宏观角度的节点(群)的运动模式。其基本步骤是:

(1)任意节点  $n_k$  记录和其它节点  $n_j$  间链路的容量函数  $c_{kj}(t)$ 。但是,一个连续时间的函数不便于存储,并且,底层协议也很难提供一个连续的  $c_{kj}(t)$ 。本文参考文献[7],利用采样和平滑对  $c_{kj}(t)$  进行离散化处理。首先给定采样间隔  $\delta_T$  和平滑步长  $\Delta$ ,通过采样得到初始序列  $c_{kj}(t_i), i=0,1,2, \dots$ ,然后按间隔  $\Delta$  进行平滑,即在  $\Delta$  内链路连通的时间占比例大于给定参数  $\rho$  时认为该段时间链路是连通的,否则是断开的。从而得到最终的  $\{c_{kj}(t_i), 0 \leq i < T\}$ 。另外,参数  $T$  给出了序列的长度, $T$  取决于存储容量和运算复杂度的限制。

(2)节点散发自己所有的记录到  $\{c_{kj}(t_i), 0 \leq i < T, 0 \leq j < N\}$  网络中,其中  $N$  为节点总数。这个过程可以由路由信息交换协议捎带完成。这里有两个问题需要考虑,一是判断收到的  $\{c_{kj}(t_i), 0 \leq i < T\}$  序列的实时性。这一点有路由信息交换协议来保证;更重要的是 TOM<sup>2</sup>D 并不要求实时性,因为它考察的是宏观的运动模式,是一个长时段运动特性。二是不同节点散发的  $\{c_{kj}(t_i), 0 \leq i < T\}$  序列的同步性。例如节点  $n_1$  记录了 0s-100s 区间,  $n_2$  记录的是 50s-150s 区间,

那么把这两个节点记录的序列组成 3D 矩阵  $C$  后,识别出来的运动模式必然是不准确的。而设置全局同步的时钟不现实,因此 TOM<sup>2</sup>D 为每个  $\{c_{kj}(t_i), 0 \leq i < T\}$  序列引入两个参数  $t_k$  和  $\delta_{kj}$ ,  $t_k$  是节点  $n_k$  记录  $c_{kj}(t_0)$  的本地时间,  $\delta_{kj}$  则是节点  $n_j$  相对于  $n_k$  的时间漂移;利用 RTT(Round Trip Time) 可以确定  $\delta_{kj}$  误差在秒级以下,对于 DTN 环境来说这种误差可以忽略。这样,节点  $n_j$  接收到来自  $n_k$  的  $\{c_{kj}(t_i), 0 \leq i < T\}$  序列后可以由  $t_k + \delta_{kj}$  来确定该序列起始时刻对应的本机时间。这样可以计算出所有接收到的序列的起始时间的本机时间,因此可以判断序列之间的同步失调程度,从而选择丢失失调严重的序列。

(3)通过步骤(2)获得了全局的链路容量矩阵<sup>1)</sup>  $C = \{c_{ij}(t_k)\}_{N \times N \times T}$ ,即一段时间内所有链路的通-断信息。然后, TOM<sup>2</sup>D 将利用图像处理中模板运算机制对  $C$  矩阵进行处理,所不同的是  $C$  是一个 3D 矩阵。模板运算的运算量小、形式统一且易于调整。识别的具体过程在第 2.2 节阐述。

(4)把识别出来的运动模式存储在数据结构 CL 中,CL 的结构如图 1 所示。图中 ListID 唯一标识一个列表同时指示运动模式的类型, Addr<sub>0</sub> 是指当前节点(用于周期运动模式中的主节点), Length 是指该列表中的节点数目,每个节点有 Addr<sub>i</sub>, Time<sub>i</sub> 和  $C_i$  分别表示节点地址、与 Addr<sub>0</sub> 的连接时间和链路容量。CL 是 TOM<sup>2</sup>D 为路由算法提供信息的统一结构,路由算法解析各 CL,并用于路由决策。

ListID	Addr <sub>0</sub>	Length	Starttime	Addr <sub>1</sub>	Time <sub>1</sub>	C <sub>1</sub>	...
--------	-------------------	--------	-----------	-------------------	-------------------	----------------	-----

图 1 Contact List 的结构

### 2.2 运动模式识别

TOM<sup>2</sup>D 用模板运算的方式从 3D 矩阵  $C$  中提取运动模式。这里的模板运算和图像处理中使用的又有一些差别:首先,不能依赖于节点的排列顺序。在图像处理中,像素的位置是确定的,而这里却和节点的顺序有关,因此我们的模板运算必须考虑这一点。其次, TOM<sup>2</sup>D 中的模板是动态的。例如,要识别周期运动,必须首先得到周期的大小,然后确定循环具体模式,才能得到模板。最后 TOM<sup>2</sup>D 的模板的运算方式更复杂。传统的模板运算是对应元素相乘后求和,并把结果赋给中心元素。而本文中模板给出的是计算方法,从一个输入矩阵得到另一个输出矩阵;也就是说不再是原位运算。可以说 TOM<sup>2</sup>D 使用的是泛模板运算,用到了多个元素整体运算从而极大的降低了复杂度。这一节我们给出组运动模式和周期模式的识别步骤。

<sup>1)</sup>本文中类似  $C = \{c_{ij}(t_k)\}_{N \times N \times T}$  的记法,表示一个 3D 矩阵,其中  $0 \leq i < M, 0 \leq j < N, 0 \leq k < T$ ; 后文中  $\{c_{ij}(t_k)\}_{N \times T}, \{c_{ij}(t_k)\}_{M \times T}, \{c_{ij}(t_k)\}_{M \times N}$  的标记则标识给定其中的一个下标得到的 2D 矩阵,依次类推。

组模式指一群节点长时间处于很高的连通状态,形成一个直观上具有相近运动特性的一个组。组模式是最常用的一种模式,现在的关于运动模式的研究几乎都是针对组模式的;但是对组的识别还是基于邻居表,即依赖于某个时刻的状态。这对于具有间歇性连接特性的网络来说并不合适,因为间歇性连接使得邻居表变化频繁,从而组结构不稳定。在 TOM<sup>2</sup>D 中采用 3D 矩阵来标识一个时间段的连接情况,进行整体考虑,其识别过程分三步:(1)在矩阵  $C$  中,对任意节点  $n_k$  和  $n_j$ ,  $\{c_{kj}(t_i), 0 \leq i < T\}$ , 统计其中 0 的个数  $m$ , 并<sup>2)</sup>记  $M^1_{kj} = -m$ , 从而得到 2D 矩阵  $M^1_{N \times N}$ , 这个矩阵的每个元素标识一个节点对之间断开的时间的累计;(2)给定门限  $\delta_G$ , 如果  $M^1_{kj} > \delta_G$  则令  $M^1_{kj} = 1$ , 否则为 0, 就是说, 确定两个节点是否属于同一个组。(3)从矩阵  $M^1$  中提取组, 并存储到 CL 中; 第(2)步只是给出了任意节点对是否属于一个组, 但是一个组有哪些成员还需要特定的算法来提取。TOM<sup>2</sup>D 用的提取算法如下(其中  $L^1_i$  是指  $M^1$  的第  $i$  行,  $L^2_i$  是  $M^2$  的第  $i$  行):

```

Input :  $M^1$ 
Output :  $M^2$ 
Initiate :  $M^2 = \{0\}_{N \times N}$ 
for each  $i(0 \leq i < N)$ 
    for each  $j(0 \leq j < N)$ 
        if  $L^1_i \& L^1_j > 0$ 
             $L^2_i = L^1_i | L^1_j$ 
             $L^2_j = 0$ 
         $L^2_i = L^1_i$ 
end

```

其基本思想是利用“或”运算把处于同一组的节点用  $M^2$  的一行表示。然后把  $M^2$  的每一行保存到一个 CL 即可。

社区模式是扩展的组模式。它用于描述一些节点定期在一定时间段内聚合成为一个组的情形。例如校园的课堂, 只有在上课时间才会有很多的学生聚合在一起。社区模式首先要找出组形成的时刻, 即节点群什么时候聚合在一起成为一个组, 然后利用组模式进行识别。具体操作如下:(1)对矩阵  $C$  进行累计差分, 即对  $c_{kj}(t_i)$  和  $c_{kj}(t_{i+1})$ , 如果  $1 \rightarrow 0$  则置为 -1,  $0 \rightarrow 1$  则为  $M^3_{k,j,t_{i+1}} = 1$ ; 然后对连续“1”累加 1, 连续“0”累加 -1, 这样得到 3D 矩阵  $M^3$ 。(2)用模板  $M^4 = \{1\}_{N \times N \times D}$  对  $M^3$  处理, 即对于每个  $t_i(D/2 \leq i \leq N - D/2)$ , 计算

$$M^5_{t_i} = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \sum_{d=-D/2}^{D/2} (M^3_{m,n,t_i+d} \times M^4_{m,n,d})$$

得到  $M^3$  关于时间轴的“重量”(即  $M^4$  给定的范围内“1”的个数)分布, 用序列  $M^5$  存储。(3)找出  $M^5$  的极值点所对应的时刻  $\{t_0, t_1, t_2, \dots\}$ , 各时刻对应的  $C$  矩阵的一个剖面即为组形成时刻的全局拓扑, 最后利用前面提到的组模式识别

方法来提取组, 结合时间信息即得到社区模式。

周期模式描述某个节点周期的访问一系列节点的情形, 包括被访问的节点的 ID 和访问持续时间两种信息。其步骤是:(1)对矩阵  $C$  中与考察节点  $n_h$  相关的 2D 矩阵  $\{c_{hj}(t_i)\}_{1 \times N \times T}$ , (其中  $h$  是给定的,  $0 \leq j, i < N$ ) 进行累计差分得到 2D 矩阵  $M^6$ , 过程如前所述; 然后找出  $M^6$  的极值点, 把相应位置元素赋 1, 其余为 0。(2)截取一定长度  $M^7 = \{M^6_{kj}(t_i)\}_{N \times N \times D}$ , 调整  $D$  的大小使得对每一个  $l = 0, 1, \dots$ ,  $M^7$  与  $\{M^6_{kj}(t_i)\}_{N \times N}$ ,  $m \times D \leq i < m \times (D + l)$ , 按元素相乘后求得和得到序列  $M^8_l(l = 0, 1, \dots)$  的每个值都大于门限  $\delta_s$ 。(3)  $M^7$  每一列即为访问节点列表, 然后对  $M^7$  按时间维(即  $t_i$ )查找其中 1 连续出现的次数即得到与每个节点连接的持续时间。

周期模式有一个简化模式, 即摆渡模式。它不关心连接时间信息, 但是要求访问节点列表包含了网络中的所有节点。

### 3 TOM2D 的使用和仿真

TOM<sup>2</sup>D 是独立于路由协议的, 它识别出来的运动模式(由 CL 存储)为路由决策提供信息; 因此它可以适配各种路由协议。这一节我们对 DSDV 和 OLSR 进行简单的修改, 使之支持 TOM<sup>2</sup>D(分别称为 TDSDV 和 TOLSR), 并通过仿真验证 TOM<sup>2</sup>D 的有用性。

由于运动模式使用统一的数据结构 CL 存储, 因此可以向路由决策模块提供统一的接口以确定下一跳:(1)查找路由表, 如果到目的节点距离小于  $d$  跳(仿真中我们取  $d=3$ ), 则按原协议进行。(2)否则, 查看目的节点是否在 CL 中, 如果是, 则选取 CL 中一个节点(如果是周期模式则选择 Addr<sub>0</sub>)为下一跳; 否则按原协议进行。从上述步骤可以看出, TOM<sup>2</sup>D 只是增加了路由选择流程的一个支路。

仿真中共有 20 个节点, 节点  $n_0$  作周期运动依次访问  $n_1 \sim n_8$ ; 节点  $n_{10} \sim n_{14}$  是一个组; 其它节点随机移动。整个场景  $1000 \times 1000 \text{m}^2$ 。仿真开始时, 随机选取 20 个节点对作为 CBR(Constant Bit Rate)流的两端; 每个流每秒发送 10 个数据包, 数据包长 520 字节。仿真时间 500s, 结果如图 2-图 4 所示。

图 2 显示投递率(成功投递的数据包占总发送量的比例)随通信范围的增加而增加, 但 TDSDV 明显好于 DSDV, 平

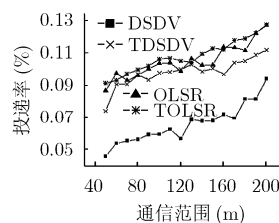


图2 投递率与通信范围的关系

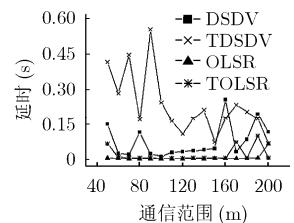


图3 延时与通信范围的关系

<sup>2)</sup>本文中如果没有特别说明, 矩阵表示法  $M^k$  中  $k=1, 2, 3, \dots$  是上标, 而不是指矩阵  $M$  的  $k$  次方。

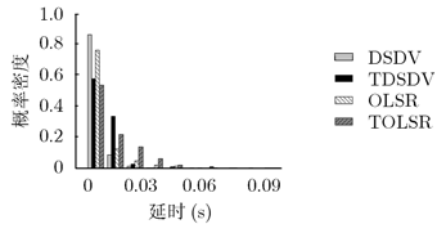


图 4 延时的分布(通信范围 100m)

均提高 49.5%；但 TOLSR 的改善相对小，这是因为 OLSR 散发 TC(Topology Control)消息，对拓扑的感知强于 DSDV。另外，整体上看，投递率都很低，因为网络分割很严重；但是当通信范围到 230m 时网络连通度极大增强，投递率会有一个跳跃的提高，达到 0.8 以上(数据跨度太大，图中没有画出)。

图 3 显示了延时。TDSDV 的延时比 DSDV 高很多，这正说明了 TDSDV 把 DSDV 中本会丢弃的数据包成功的投递了，这些数据包带来了很大的延时。就 DTN 来说，投递率是关注的重点，以延时的增加为代价是值得的。TOLSR 则不明显，这与 TOLSR 相对于 OLSR 投递率提高不明显是一致的。

图 4 是通信范围 100m 时所有成功投递的数据包所经历的延时分布情况，可以看出 TDSDV 和 TOLSR 有更长的拖尾，即成功投递了更多的高延时数据包——这正是 TOM<sup>2</sup>D 的目的。

#### 4 结论

本文提出了一种用于 DTN 的节点运动模式识别机制 TOM<sup>2</sup>D。它是一种基于泛模板运算的统一框架，能够为路由协议提供路由决策信息。本文详细介绍了组模式和周期模式的识别步骤，并把它们用于 DSDV 和 OLSR。仿真结果表明 TOM<sup>2</sup>D 能够“挽救”那些因为网络分割而被丢弃的数据包，虽然增加了平均延时，但对投递率的改善很明显。

但是 TOM<sup>2</sup>D 还有进一步研究的可能，如可以引入更多运动模式，让运动模式在路由决策中的更全面的应用等。

#### 参考文献

- [1] 于宏毅等. 无线移动自组织网[M]. 北京: 人民邮电出版社, 2005, 第 6-8 章.
- [2] Wang K H and Li B. Group Mobility and Partition Prediction in Wireless Ad-Hoc Networks, in Proceedings of IEEE International Conference on Communications (ICC), NYC, NY, April 2002. Vol.2. 1017-1021.
- [3] Leguay J, Friedman T, and Conan V. DTN routing in a mobility pattern space. In Proceeding of the 2005 ACM SIGCOMM Workshop on Delay-Tolerant Networking Philadelphia, Pennsylvania, USA, August 26-26, 2005: 276-283.
- [4] Ghosh J, Beal M J, Ngo H Q, and Qiao C. On profiling mobility and predicting locations of wireless users. In Proceedings of the Second international Workshop on Multi-Hop Ad Hoc Networks: From theory To Reality Florence, Italy, May, 2006: 55-62.
- [5] Musolesi M and Mascolo C. A community based mobility model for ad hoc network research. In Proceedings of the Second international Workshop on Multi-Hop Ad Hoc Networks: From theory To Reality Florence, Italy, May, 2006: 31-38.
- [6] Ghosh J, Qiao C, And Philip S J. Sociological orbit aware routing in manet. Tech Report, CSE Dept. of University at Buffalo, Buffalo, NY, TR-2005-09(March 2005).
- [7] Merugu S, Ammar M, and Zegura E. Routing in space and time in networks with predictable mobility, Georgia Institute of Technology, Technical Report GIT-CC-04-7, 2004.

周晓波: 男, 1981年生, 博士生, 研究方向为宽带无线网络技术和DTN.

张 幸: 女, 1984年生, 博士生, 研究方向为ICN (Intermittent Connected Network).

彭 敏: 男, 1984年生, 博士生, 研究方向为移动Ad hoc.

卢汉成: 男, 1977年生, 博士生, 讲师, 研究方向为下一代互联网技术和无线网络技术.

洪佩琳: 女, 1960年生, 教授, 研究方向为信息通信网技术.